ENVS 193DS Homework 1

Madison Conrad

04-15-2025

Homework set up

```
# insert code to load in tidyverse here

library("tidyverse")

## — Attaching core tidyverse packages ———————————————— tidyverse
2.0.0 —
## ✔ dplyr      1.1.4      ✔ readr      2.1.5
## ✔ forcats    1.0.0      ✔ stringr    1.5.1
## ✔ ggplot2    3.5.1      ✔ tibble     3.2.1
## ✔ lubridate  1.9.4      ✔ tidyr      1.3.1
## ✔ purrr      1.0.4
## — Conflicts ———————————————————————————————
tidyverse_conflicts() —
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force
all conflicts to become errors

# insert code to read in data (glacial_volume_loss_copy.csv) and save as
object called `glaciers` here

glaciers <- read.csv("glacial_volume_loss_copy.csv", stringsAsFactors =
FALSE)
```

Problem 1. Measures of central tendency and data spread

a.

[This data set is continuous/numerical data because it can be measured and represents the length in centimeters of the salamanders]

b.

```
# insert code to calculate the mean here

mean(4.6,4.4,6.2,5.2,3.7,6.0,3.9,4.6,2.7)

## [1] 4.6
```

[The mean is 4.6 cm]

c.

```
# insert code to calculate the variance here

x<-c(4.6,4.4,6.2,5.2,3.7,6.0,3.9,4.6,2.7)
sample_variance<-var(x)
```

[The sample variance is 1.2 cm]

d.

```
# insert code to calculate the sd here

lengths_cm<-c(4.6,4.4,6.2,5.2,3.7,6.0,3.9,4.6,2.7)
sample_sd<-sd(lengths_cm)
```

[The standard deviation is 1.1 cm]

Problem 2. Visualizing data

a.

[The data glacial_volume_loss_copy.csv is different from glacial_volume_loss.csv because it is just the raw data organized in a table, whereas the other includes this raw data table but also explanation of what each column means, what units were used, and information about who collected the data and where it came from]

b.

Who

[Who collected and processed the data?] [The data was collected and processed by Dyurgerov, M. Edited by W. Meier and R. Armstrong.]

What

[What project were they collected under?] [This data set was collected under the NSIDC: National Snow and Ice Data Center to obtain Glacier Mass Balance and Regime Measurements and Analysis from 1945 to 2003.]

Why

[Why were the data collected?] [The data was collected to observe glacier regime parameters such as cumulative glacial volume change.]

Where

[Where were the data collected and processed?] [This data was collected in each continent and was processed and consolidated in Boulder, Colorado, USA.]

When

[When were the data collected?] [The data was collected between 1945 and 2003.]

How

[How do I access or the data?] [The data can be accessed from the Institute of Arctic and Alpine Research (INSTAAR).]

C.

```r
# insert code to create the histogram

library(ggplot2)

ggplot(data = glaciers, #start w/ glacial data
       aes(x = Annu_sea_rise)) + #want sea level rise in mm in x-axis,
already does count on y-axis
  geom_histogram(bins= 10,
                 fill = "turquoise",
                 color = "black") +
  labs(x = " Annual Sea Level Rise (mm)", #name x-axis
       y = "Frequency",  #name y-axis
       title = "Histogram of Annual Sea Level Rise") #give & name title
```

Histogram of Annual Sea Level Rise
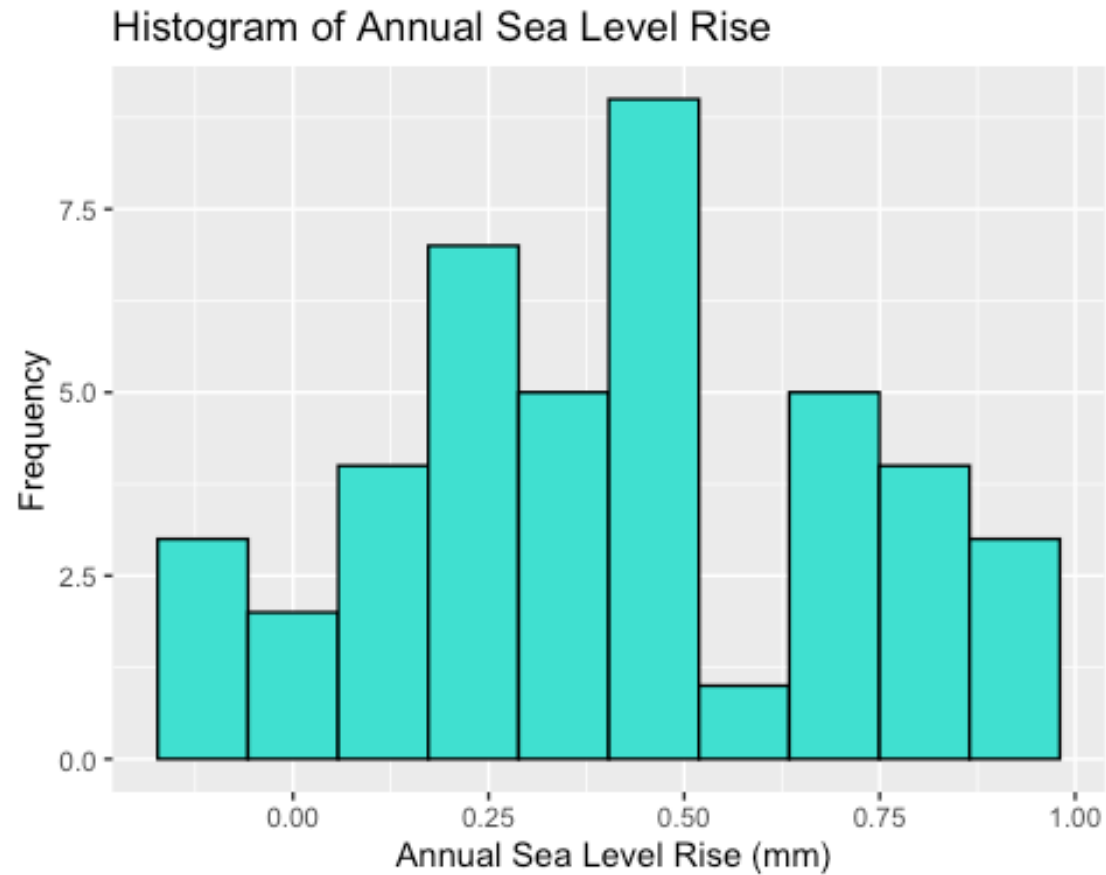
d.

```
# insert code to create the scatterplot

ggplot(data = glaciers, #start w/ glacial data
       aes(x = Year, #want years on x-axis
           y = Cumu_sea_rise )) + #want cumulative sea level rise in mm on y-
axis
  geom_point(color = "blue") +
  geom_line(color = "black") +
  labs(x = "Year", #name x-axis
       y = "Cumulative Sea Level Rise (mm)", #name y-axis
       title = "Cumulative Sea Level Rise by Year") +#give and name title
  theme_minimal()
```

Cumulative Sea Level Rise by Year
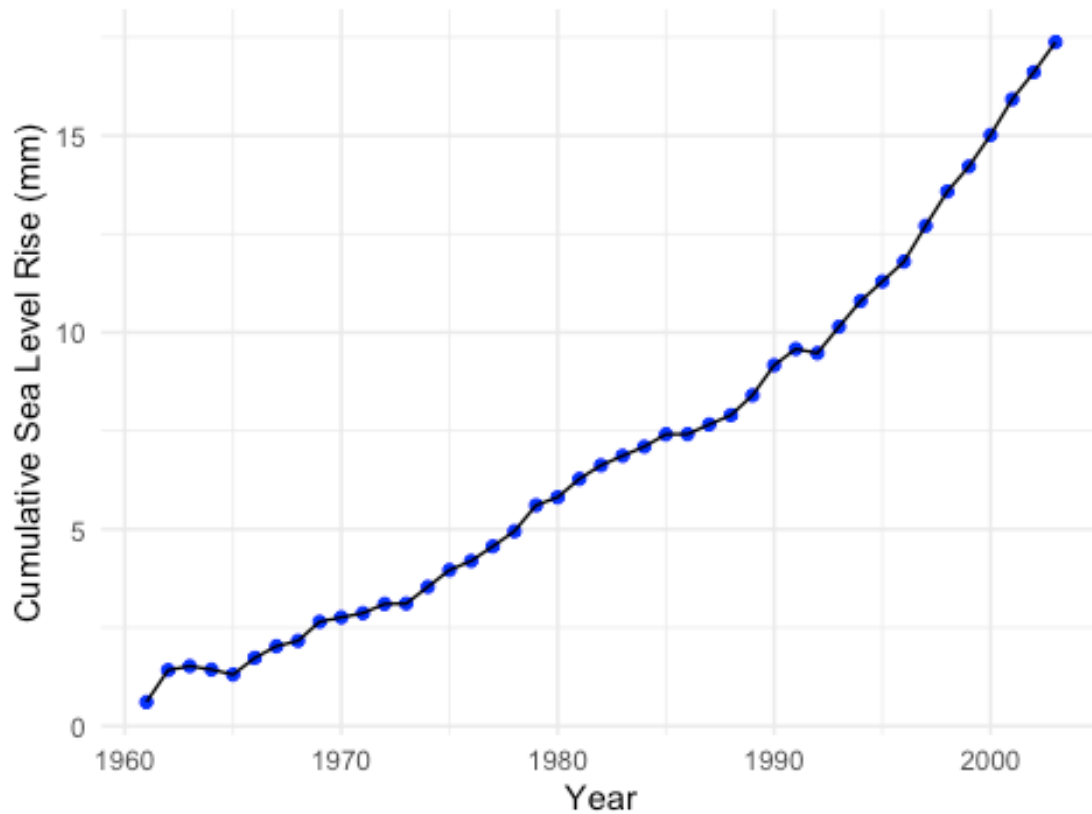
Part 3. Personal data

Idea 1. - time spend on morning walk or time

a.

[Does listening to music while stuyding or doing work affect how productive I feel?]

b.

[My response variable would be a self-rated productivity level on a scale of 1 (not productive at all) to 10 (very productive)]

c.

[My predictor variable would be whether music was playing during the study session (yes/no) which is a numerical/continuous variable]

d.

[To measure the time-related variables I would take note of the date and the start time of the study session]

e.

[Four additional variables would be the type of music I'm listening to, the length of the study session, the class or category of work I'm doing, and how distracted I am while working]

f.

[Type of music (genre): categorical Length of study session (hours): numerical/continuous Class/category working on (class or job name): categorical Distractions (scale of 1-10): numerical/continuous]

g.

[The data for these variables would be self-reported and would be supported by my Spotify and Youtube history for the music genre and length of study session]

h.

[I would record the data when I am done working]

i.

Date	Start Time	Productivity Level (1-10)	Type of Music	Length of Study Session (hours)	Class/Job	Distractions (1-10)
4-14-25	9:56 PM	8	Classical	1 hour	ED 197	4
4-15-25	7:34 PM	5	Classical	2 hours	ED 197, ENV S 193DS	9
4-16-25	9:35 PM	7	Classical	1 hour	ENV S 193DS	6

Idea 2.

a.

[Is there a relationship between biking to campus and my mood afterwards?]

b.

[My response variable would be my mood level after biking to campus ranked on a scale of 1 (awful/very bad) to 10 (awesomely great). This is a numerical variable (ordinal)]

c.

[My predictor variable would be the time I spend biking to campus measured in minutes. This is a numerical variable]

d.

[To measure the time-related variables I would take note of the date and the start and end time of my bike ride]

e.

[Four additional variables could be the weather conditions, the temperature outside , the amount of sleep I got the night before, and if I'm listening to music during my bike ride]

f.

[Weather conditions (sunny, windy, cloudy, rainy, other): categorical Temperature(Fahrenheit): numerical/continuous Amount of sleep (hours): numerical/continuous Music (yes/no): categorical]

g.

[The data for these variables would be self reported and would be supported by the weather app for the temperature and weather conditions and the sleep tracker on my phone]

h.

[I would record the data immediately after arriving on campus and at home]

i.

Date	Start Time	End Time	Mood (1-10)	Weather	Temperature (degrees F)	Amount of Sleep	Music?
4-14-25	12:21 PM	12:29 PM	7	sunny	61	7 hrs 31 min	yes
4-15-25	12:47 PM	12:58 PM	8	cloudy	60	8 hrs 22 min	yes
4-16-25	8:51 AM	8:59 AM	7	cloudy	60	6 hrs 45 min	yes

Problem 4. Setting up statistical critique

Remember to fill out the Google form with DOIs!

Paper 1

[Ambrosino, C.M., & Rivera, M.A.J. 2022. "A longitudinal analysis of developing marine science identity in a place-based, undergraduate research experience." IJ STEM Ed 9:70. https://doi.org/10.1186/s40594-022-00386-4]

Paper 2

[Zhang, C. et.al. 2023. "Marine mammal morphometrics: 3D modeling and estimation validation." Frontier Marine Science 10:2023. https://doi.org/10.3389/fmars.2023.1105629]

Paper 3

[Beninger, P.G. et.al. 2012. "Strengthening statistical usage in marine ecology." Journal of Experimental Marine Biology and Ecology 426:2012.https://doi.org/10.1016/j.jembe.2012.05.020]